.

MINE DISCRIMINATION USING MULTISPECTRAL IMAGERY WITH FEEDFORWARD NEURAL NETWORKS

Taher Daud, Tuan Duong, Harry Langenbacher, Helen Tsu, and Anil Thakoor
Center for Space Microelectronics Technology
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California91109

1. ABSTRACT

As a precursor to hardware implementation, wc have performed mine detection functions in simulation on a polarimetric hyperspectral imaging (PI II) dataset collected by using a technique of acousto-optic tunable filter (AOTF) camera. In principle, PHI data of an image containing objects of interest in a cluttered background provide significant information about the. objects, including their relative sizes, shapes, orientation, and other characteristics such as light reflectance and polarization signatures based on their material properties. The present study was, however, limited only to the direct spectral data for the object of interest.

A feedforward artificial neural network (ANN) architecture was programmed to recognize predefined spectral "templates" by using a well-known, hardware implementable inner-product matching scheme. This scheme is particularly suited to the problem of spectral discrimination where the spectra to be examined or the objects to be discriminated are uncorrelated, as in the present case. In this paper, wc describe the ANN architecture and discuss its hardware implementation issues. In addition, we provide the results of our simulation study performed along with the suitable preprocessing steps with various window sizes from 1 xl to 50x50 pixels, leading to the unambiguous detection of the positions of mines in our test runs without false slam-ts.

2. INTRODUCTI ON

Pattern recognition is an important class of problems that requires intensive data processing. The pattern recognition function is fairly complex, specially when the input is noisy or incomplete. In addition, there generally may not exist a formalism of the algorithmic input-output relationships. Some of the examples of such "mapping" application are identification of minerals using multispectral data, target classification based on infrared (lR), visual, or acoustic data, forensic analysis of fingerprints and signature matching, etc. Where the categories involved in a given scene arc well correlated, a feedforward neural network is capable of "learning" the input-output relationships to perform as a classifier. However, in many object recognition and classification functionalities, one class differs from others with no correlation, and learning algorithms may not be suitable or maybe wasteful because of time required for "learning". In these cases, template matching schemes of feedforward neural networks arc more suitable, and when implemented in parallel hardware, high speed can also be achieved.

Mine detection using side-scan sonar data with cluttered background has been shown to perform well with a trained feedforward neural network[1]. However, mineral identifications from multispectral data has been shown using template/pattern matching algorithm [2]. Pattern matching involves a suitable comparison of the inner product of the input pattern elements with the respective elements of the stored patterns or templates, and selection of the best match. This method is analogous to a binary pattern recall method with minimum Hamming distance as a measure of similarity between an unknown input pattern and one of several stored patterns [3]. Template matching is particularly suitable when the data is in the form of spectral response and hence can be converted as a line vector for ease of matching with stored standard and predetermined signatures.

Polarimetric hyperspectral imaging (PHI) is an effective remote sensing technique and uses a visible/near infrared acousto-optic tunable filter (AOTF) prototype system. The data gathering is done in a range of wavelengths that depends on the detectors used, normally in the visible and/or near infrared spectrum (0.48 to 0.75 μm) [4]. This technology offers promise of differentiating between natural and man-made objects, and has been shown to offer discrete signature spectra for vegetation, minerals, and atmosphere depending on their surface texture, ingredients, etc. which are particularly sensitive to polarized light. Target detection capabilities are enhanced because of such perceptive data collection and analysis, and objects with camouflage or in cluttered background have a better chance of detection with such hyperspectral data analysis capability

We report on a feedforward neural network pattern matching algorithm modeled after an associative content addressable memory which is an analog equivalent of the binary inner product scheme of Ref. [3]. This scheme was tested for detection of mine-like objects (MLO) where multispectral data across a terrain using an acousto-optic tunable filter (AOTF) has been obtained. Two types of objects, one plastic and the other steel, were used for this classification scheme along with background scenery, other manmade objects and clutter in the scene.

We briefly describe the neural network pattern matching algorithm and the need to normalize both the stored and the input patterns. A short description of the architecture is given which lends itself to VLSI implementation if needed for obtaining high speed. The problem of discrimination of MLO using AOTF data is described. This is followed by our experimental results and discussions.

3. PATTERN MATCHING ALGORITHM

The pattern matching algorithm is based on a feedforward neural network model with an input layer, an intermediate processing layer and an output layer. The input layer stores the preselected patterns and after receiving an input pattern, generates inner products of the input pattern with all the stored patterns (in a fully parallel fashion, and hence at high speed, when implemented in hardware). The intermediate layer is designed to perform the functions of finding the best match (in some cases, n best matches, if desired), and providing label for the selected pattern. It is also feasible to provide an output layer in hardware that reproduces the winning pattern from storage as the required output pattern.

Let us assume that a set of m patterns Y^{j}, each as a multi-element vector,

$$Y^j = (y_1^j, y_2^j, y_3^j, \ldots, y_n^j) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

where j refers to the jth pattern, is stored and each stored pattern is compared with an incoming pattern

$$X = (x_1, x_2, x_3, \ldots, x_n) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Then, a correlation of comparison is given by the respective inner products of (1) with (2) as,

$$C_j = \sum_{i=1}^{n} x_i y_i^j \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

As there are m stored patterns, there will be m different values of C to be compared. When the incoming pattern is closest to one of the stored patterns, that inner product will be highest in value. Once discerned, a classifier label of the stored pattern associated with maximum correlation C_{max} can be provided. It may be noted that an exact match is not necessary. By the same token, an incomplete input pattern is equally valid provided it elicits the response C_{max} at the right label. However, one needs to consider the following two points: one, it is important that the pattern vectors be of equal lengths; and second, each vector should be normalized (say, to a common base) so that large- and/or small-valued elements in some of the patterns do not skew the results [2].

4. PATTERN NORMALIZATION

We have investigated two possible procedures of pattern normalization. The first normalizes each pattern as a whole so that all the patterns are converted to a common base. The second procedure normalizes each individual element of the patterns to a common base. These are explained below.

4.1 Normalization Procedure I

Consider a pattern vector

$$z^j = (z_1^j, z_2^j, z_3^j, \ldots, z_n^j) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

We can normalize this vector so that the normalized vector

$$y^j = (az_1^j, az_2^j, az_3^j, \ldots, az_n^j) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

A perfect match between a stored and an input vector is obtained when their inner product is a maximum. Thus,

$$\sum_{i=1}^{n} (y_i^j)^2 = C_{max} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

Therefore,

$$\sum_{i=1}^{n} (y_i^j)^2 = \sum_{i=1}^{n} (az_i^j)^2 = C_{max} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

giving the normalization constant for that vector,

$$a = \sqrt{\frac{C_{max}}{\sum_{i=1}^{n} (z_i^j)^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (8)$$

The correlation of the jth normalized stored vector az^j with an arbitrary normalized input vector bx is,

$$C_j = \sum_i ab\, x_i\, z_i^j \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (9)$$

$$C_j = C_{max}\left\{\frac{\sum_i x_i z_i^j}{\sqrt{\sum_i (x_i)^2 \sum_i (z_i^j)^2}}\right\} \quad \ldots\ldots\ldots\ldots \quad (10)$$

When the input pattern vector is identical to one of the stored pattern, the term in brackets is unity. For a mismatch, it can be shown that the term in brackets will be less than unity and is a measure of goodness of correlation.

4.2 Normalization Procedure II

In this procedure, for each element of a pattern there is added a complementary element. This procedure thus converts an n-element pattern to a 2n-element long pattern. Consider a pattern vector r with n elements, $r_1, r_2, \ldots, r_i, \ldots, r_n$. Here, each element r_i is associated with a complementary element S_i such that $(r_i)^2 + (s_i)^2 = x$; where x is the normalizing constant. Here again it can be shown that for two normalized patterns (r,s; and z, y), $C_{rz} = \sum(r_i z_i + s_i y_i)$ is always less than C_{max} (= n x) unless the two patterns arc identical, and then $C_{rz} = C_{max}$.

Normalization procedure I requires computation on the whole pattern whereas procedure II deals with individual elements of the pattern at a time. Therefore, the procedure 11 is more amenable to hardware implementati on. Further, it ensures that even an element with a low value (say, zero) is given proper weighting because its complement is then highest (= x). The disadvantage, of course, is that the pattern vector is expanded to twice the original size. It is also possible to combine the two procedures if the problem at hand warrants that. Wc have adopted procedure 11 for normalization of vector elements during this test of the AOTF data for classification.

5. ARCHITECTURE

The algorithm coding for detection of MLO was done in simulation. However, wc describe the equivalent architecture which can be implemented in our analog VLSI hardware [5]. The input layer consists of a matrix of synapses and each synapse has a set of digitally controlled latches for storage of weights, Wij, and a multiplying digital to analog converter (MDAC) for obtaining the "input-W_{ij}" product [6]. Each element of input pattern is presented as a voltage and is first converted as an equivalent current. It is then multiplied with the stored weights and generates the product as a current signal. These currents along a row for all the elements of a single vector arc summed along that wire as the inner product of the input pattern elements with the elements of the stored pattern along that row. Thus, for an n-element (i = 1 to n) pattern scheme, the inner product of the input x with the jth stored vector,

$$C_j = \sum_i W_{ij} x_i \qquad \ldots\ldots\ldots\ldots\ldots\ldots \qquad (11)$$

would be available along the jth row. The architecture is shown schematically in Figure 1. No learning is involved because the patterns to be recognized against arc known and accordingly prestored. The second layer is the winner-take-all which compares the inner products C_j along individual rows to dctcn-nine C_{max} and provide the label of that row. Knowing the row label, the discrimination of the input pattern is thus obtained [2].

6. AOTF DATA ANI) SETUP

MLOS arc commonly small-size objects measuring around 30-40 cm. across. They may be metallic or plastic, and painted with color or in their nascent condition. When they arc placed in outdoor environment, light scattering and reflection properties of the surrounding affect their observable spectra duc to "spectral mixing" [6]. The problem may become even more complex for smaller targets and those that arc partially occluded by the surroundings.

For our experiment, the AOTF data was taken with two MLO types (one metallic and the other plastic) in full view for obtaining the spectral information that would be used as the reference patterns. Figure 2 shows a gray-scale image of the area where the two MLOS were placed. In addition, there were other objects of different shapes and sizes along with cement structures in a vegetated environment. Even though the AOTF data was available with both direct and polarized light signatures, only the direct spectrum was used for this short duration of the experiment, with some consequent degradation in the discrimination results.

The details of the data collection using an AOTF camera arc given elsewhere [6]. The data consisted of a multi-dimensional intensity spectra set, where each pixel on the terrain had a two-dimensional spectral signature of wavelength vs. intensity for the direct and a similar set for the polarized light. Each spectrum had 32 wavelengths with corresponding intensities.

In addition to the normalization of the data (giving 64 elements/pattern), an averaging function was used because each MLO was much larger than one pixel and had varying textures and surface reflection properties along their surface areas which led to differences in reflected spectrum from pixel to pixel on the same MLO. The pixel averaging dots tend to 1OSC some of the useful feature information available in the image; however, as a first order, it reduces the effects of variation in illumination and shadows from the spectral signature assigned to the particular object. A number of experiments were done to classify the mines by averaging the spectrum intensities for a (n x n) pixel window around each pixel within the objects, where n was varied as 1, 3, 5, 10, 20, 25, 30, and 50. Stored patterns representing metallic and plastic MLOS were thus obtained. In our exercise, except for the two distinct MLO signatures all other object spectra were classified as "Other" thus giving a total of three discrimination patterns. A limited runs were made by considering tree and ground as two additional classes. Each pattern had 64 elements and a total of at least three different classes.

7. RESULTS AND DISCUSSIONS

With the stored images to compare with, the input vectors at each pixel were used in the inner product scheme. The results of the experiments with the different averaging window sizes for the stored patterns is shown as a set of gray-scale images in Figure 3.

The winner with the largest inner product was obtained by using a variable threshold value. The procedure resulted in an unambiguous detection of the position of the mines in all the test runs, with virtually no false alarms anywhere in the image. As expected, the position of the MLOS were marked correctly without being able to mark the exact edges/contours or shapes. This may partly be due to the somewhat arbitrary and unoptimized selection of window sizes.

The discrimination of mines was a very sensitive function of the selection of threshold value. Once a comet value was chosen, the false alarms were avoided. However, there was no way of a priori determining the exact threshold value. However, in general, the procedure worked to distinguish the MLO correctly. It was expected that additionally use of polarized signatures and combination of both the normalization procedures I and II would improve the discrimination process further.

8. ACKNOWLEDGMENTS

The authors wish to thank Drs. H.R. Suiter and A. C. Dubey of NSWC/CSS for useful discussions. The research described in this paper was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology, and was jointly sponsored by the Naval Surface Warfare Center, the Ballistic Missile Defense Organization, and the National Aeronautics and Space Administration.

9. REFERENCES

[1] T. Daud, S. Eberhardt, T.X Brown, M. Tran, A.P. Thakoor, M. Dzwonczyk, M. Busa, M. LeBlanc, and T. Sims, "Neural networks building-block chips for classification and detection applications," Fifth International Conference on Neural Networks and their Applications, Neuro Nimes, Nov. 1992, pp. 565-575.

[2] T. Daud, R. Tawel, H. Langenbacher, S.P. Eberhardt, and A.P. Thakoor, "Analog parallel processor hardware for high speed pattern recognition," Proc. of the SPIE Conf. on Visual Communications and Image Processing, vol. 1360, Ed: M. Kunt; 1990, Lausanne, Switzerland, pp. 359-370.

[3] T. Kohonen, Content Addressable Memories, Springer-Verlag, Berlin, Germany, 1980.

[4]. L-J. Cheng, T-H. Chao, M. Dowdy, C. LaBar, C. Mahoney, G. Reyes, and K. Bergman, "Multispectral imaging systems using acousto-optic tunable filter," Proc. SPIE Conf. Infrared and Millimeter Wave Engineering, vol. 1874, 1993, pp. 224-,

[5] T. Daud, T. Duong, M. Tran, H. Langenbacher, and A. Thakoor, "High resolution synaptic weights and hardware-in-the-loop learning," Proc. of the SPIE Photonics '95 Conf. vol. 2424, Feb. 1995, (to be published).

[6] A. Moopenn, T. Duong, and A.P. Thakoor, "Digital-analog hybrid synapse chips for electronic neural networks," Proc. Advances In: Neural Information Processing Systems-2; Ed: D. S. Touretzky; Publishers: Morgan Kaufmann, 1990, pp. 769-776.

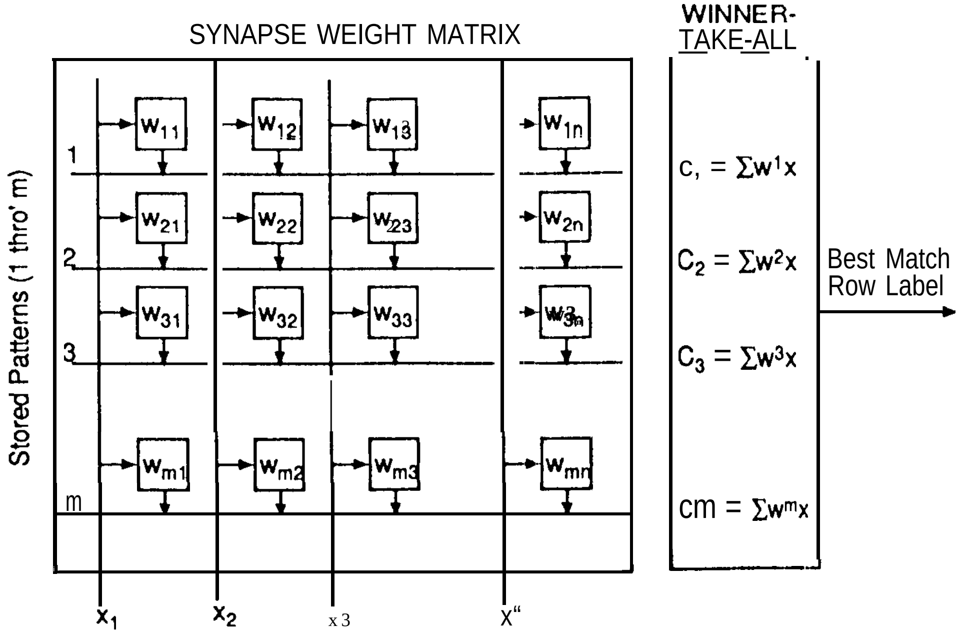
[7] L-J. Cheng and G. Reyes, "AOTF polarimetric hyperspectral imaging for mine detection," Proc. of the SPIE Conf. on AeroSense, April 17-21, 1995 (to be published).

Figures:

Figure 1. A schematic diagram of a pattern matching architecture with inner product scheme and a winner-take-all for providing the label of the winning pattern.

Figure 2. A gray scale rendition of the scene with two mine-like objects for classification, one of steel and the other of plastic. The scene also contains vegetation, other man-made objects, etc.

Figure 3. A set of gray scale pictures of the same scene where dark black and white areas are the location of MLOS and the remaining areas are gray and classified as "other". The classification is provided by using different pixel data averaging of the stored patterns: a) 1x1; b) 3x3; c) 5x5: d) 10x10; e) 20x20; f) 25x25; g) 30x30; and h) 50x50. In some cases tree (near the top of the picture) and ground (near the bottom of the picture) are also classified. Stored patterns with 10x 10, to 30x30 pixel averaging seem to classify better than with other pixel averaging window sizes.



SYNAPSE WEIGHT MATRIX
WINNER-TAKE-ALL
Stored Patterns (1 thro' m)
1
2
3
m
w_{11} w_{12} w_{13} w_{1n}
w_{21} w_{22} w_{23} w_{2n}
w_{31} w_{32} w_{33} w_{3n}
w_{m1} w_{m2} w_{m3} w_{mn}
$c_1 = \sum w^1 x$
$C_2 = \sum w^2 x$
$C_3 = \sum w^3 x$
$cm = \sum w^m x$
Best Match
Row Label
x_1 x_2 x_3 x_n
Input Pattern

Fig. 1

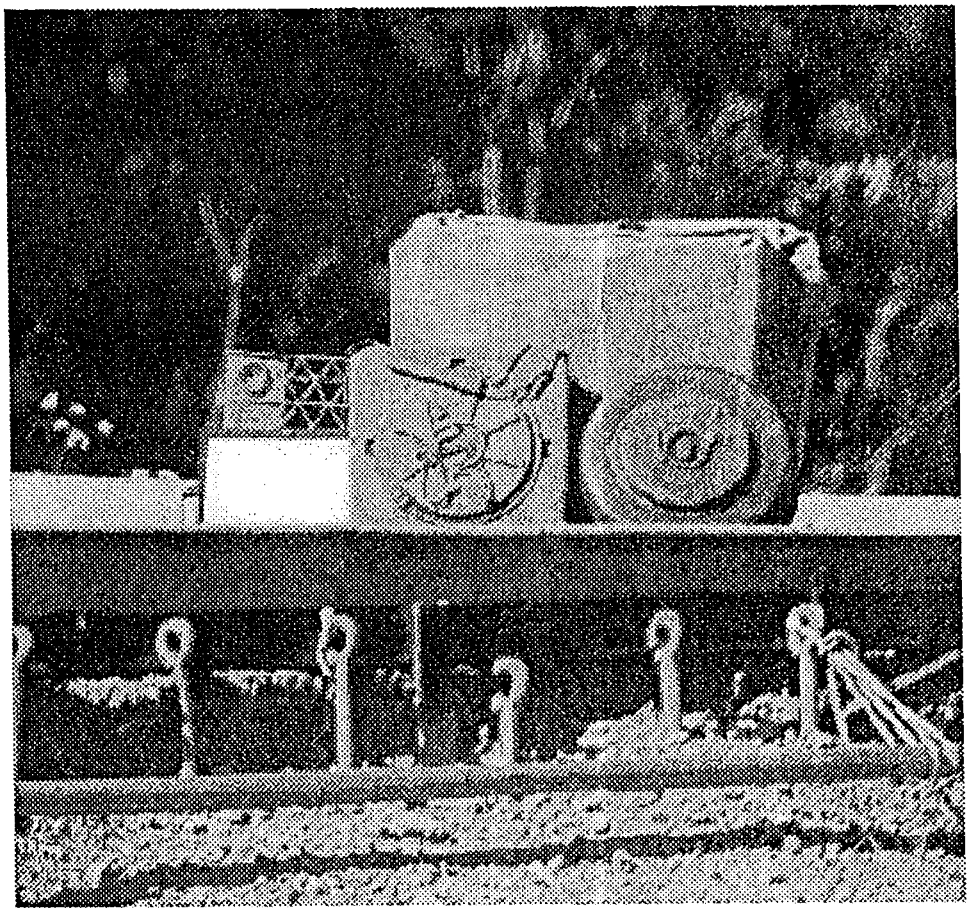
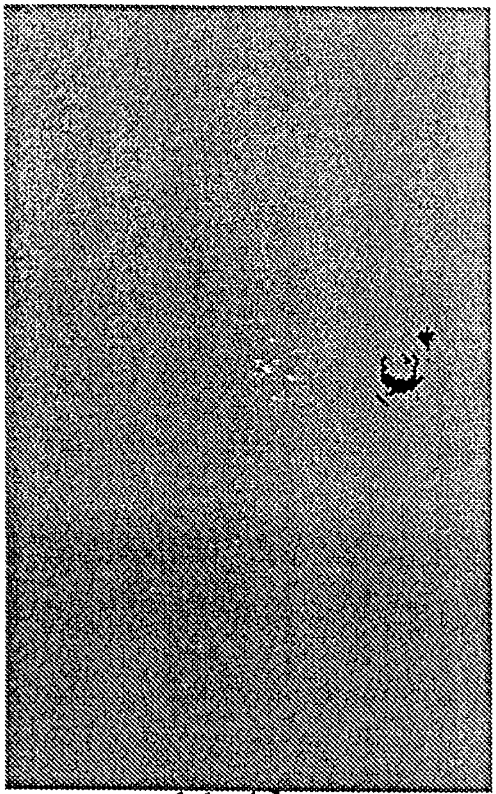
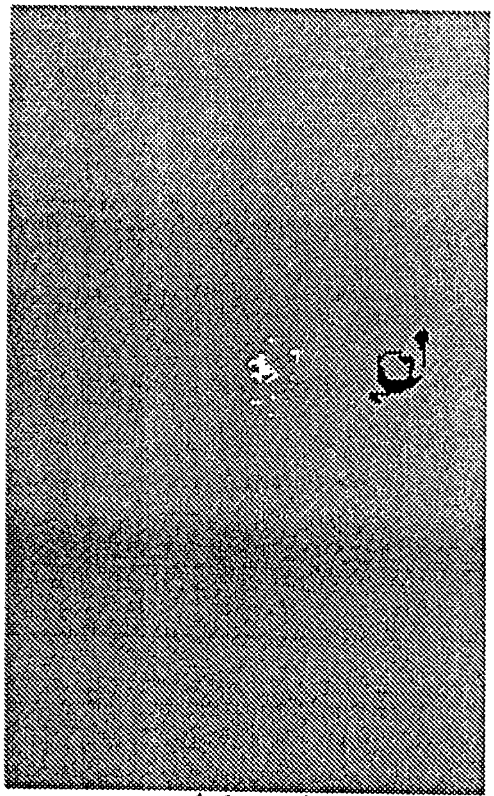


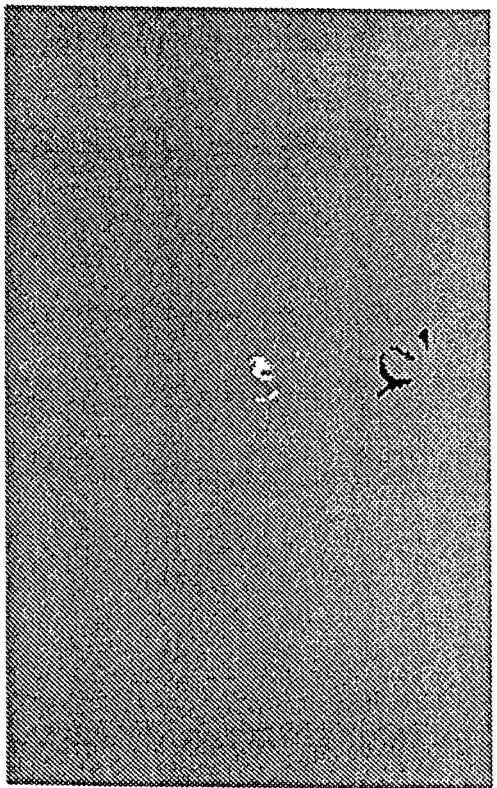
Fig. 2



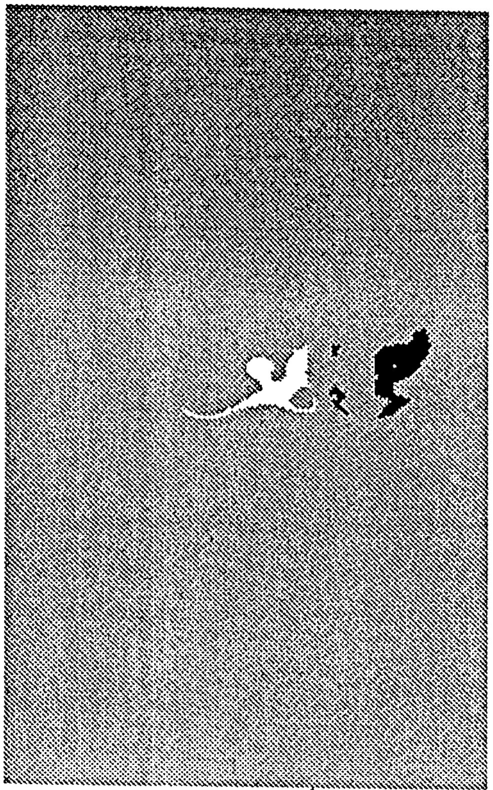
(1×1)



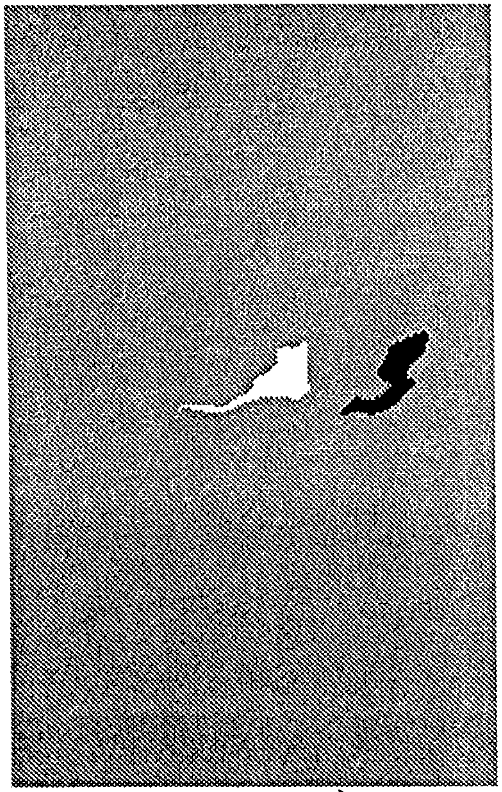
(3×3)



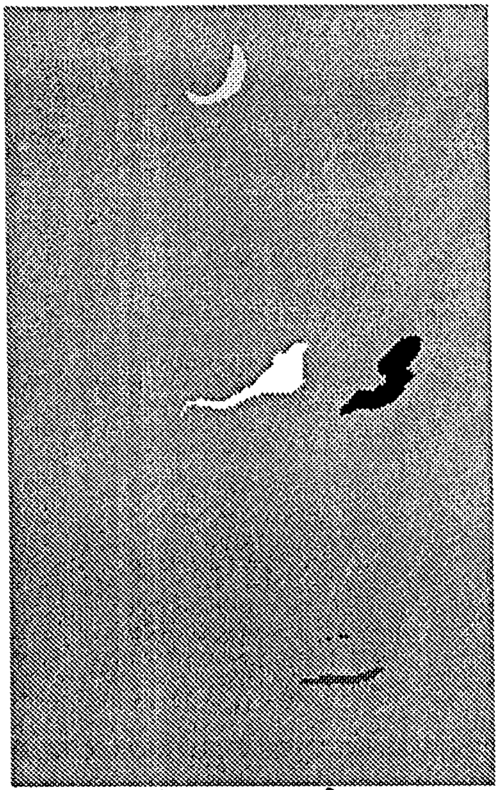
(5×5)



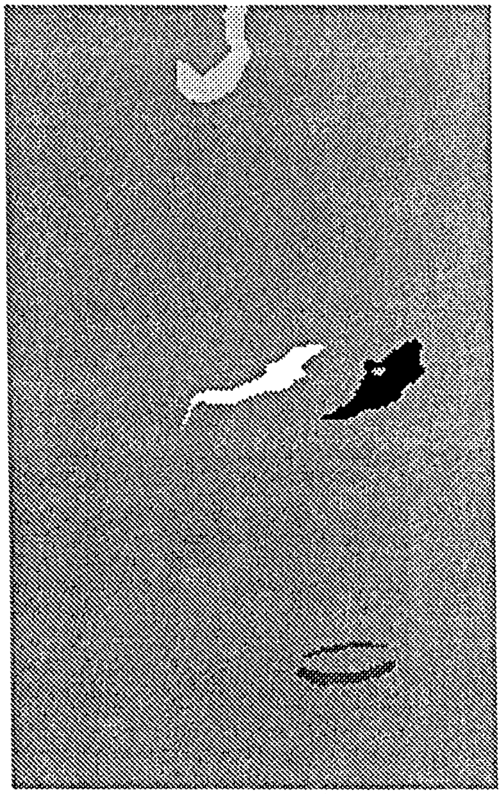
(10×10)



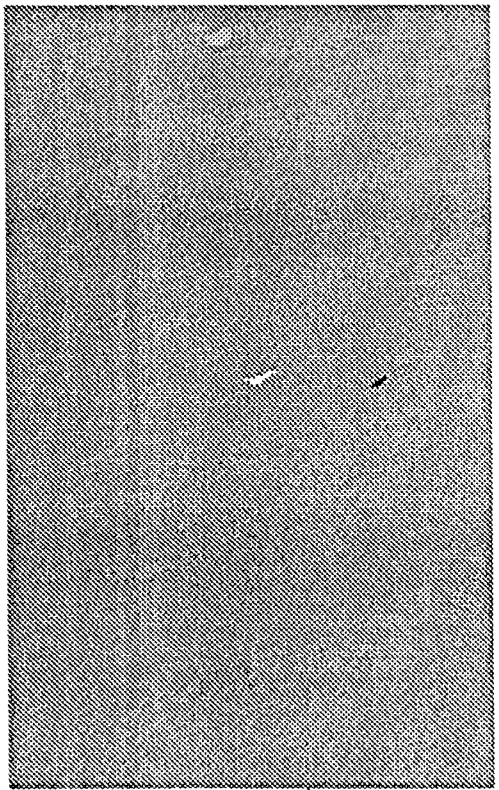
(20×20)



(25 × 25)



(30 × 30)



(50 × 50)

Fig. 3 (a-h)